Opportunistic Spectrum Access for Cognitive Radio in the Presence of Reactive Primary Users

Yue Ling Che†, Rui Zhang‡, and Yi Gong†
†School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore
Email: {chey0004, eygong}@ntu.edu.sg
‡Institute for Infocomm Research, A*STAR, Singapore, and ECE Department, National University of Singapore
Email: rzhang@i2r.a-star.edu.sg

Abstract—Opportunistic spectrum access (OSA) is a key technique for the secondary user (SU) in a Cognitive Radio network to transmit over the “spectrum holes” unoccupied by the primary user (PU). Most existing work on the design of OSA has assumed a non-reactive (NR) PU model, i.e., the PU transmission on-off status is independent of the SU access policy, which may not be practical. In this paper, we propose a new Reactive Primary User (RPU) model for the study of OSA, where the PU’s access probability over a particular channel is related to the SU’s past access history. We model the channel occupancy of the RPU as a 4-state Markov chain, as opposed to the conventional 2-state (on/off) counterpart, where the expanded state space and state transition probabilities are used to model the reactions of the PU subject to the SU transmit collision. Under this model, we formulate the optimal OSA design for the SU’s throughput maximization as a finite-horizon partially observable Markov decision process (POMDP) problem, subject to a conditional collision probability constraint for protecting the PU. Because of the high complexity of the proposed problem, we further propose a separation principle to obtain the optimal policy for the SU with implementable complexity. Numerical results show the new tradeoff between the SU’s and the PU’s throughput under the RPU model, as compared to the conventional NR PU model.

I. INTRODUCTION

Cognitive radio (CR) has emerged as a promising technology to resolve the spectrum shortage problem. Because of the competing goals of minimizing the interference to primary users (PUs) and maximizing the throughput for secondary users (SUs), opportunistic spectrum access (OSA) for CR has attracted a great deal of research attention. Among others, [1]-[3] have assumed time-slotted transmission for both SUs and PUs and studied the optimal SU channel access policies by formulating the OSA design in a partially observable Markov decision process (POMDP) framework. In [2], Chen et al. proposed the conditional collision probability constraint for the SU as a method to protect the PU, and developed a separation principle to reduce the complexity for solving the corresponding POMDP problem. OSA design for unslotted transmission is also extensively studied. For example, [4] has considered continuous-time Markov channel models via a similar POMDP approach. Also, [5] developed OSA designs by incorporating the PU traffic patterns and the packet collisions with SU.

However, most existing work on OSA, including the above prior ones, has assumed a non-reactive (NR) PU transmission model, where the PU’s on-off status over a particular channel is independent of the SU access policy. Although appealing to simplify the analysis, such assumption is usually impractical since today’s primary systems are mostly intelligent enough to adapt transmission upon receiving collision or interference. For example, PUs may increase transmit power to compensate the link loss caused by interference, or they may reduce the probability of accessing a channel when collision occurs in a carrier-sense-multiple-access (CSMA) based primary system. We refer to such PUs as reactive primary users (RPUs), to differ from their non-reactive counterparts. It is worth noting that there has been some recent work on CR considering reactive PUs. For example, the PU transmit adaptation such as power and/or rate control upon receiving a probing interference signal from the SU has been exploited in [6] and [7], to design more efficient SU interference control to the PU. A CSMA-based RPU model was proposed in [8] to investigate the performances of different SU access policies. However, [8] did not address the design of optimal access policy.

In this paper, we focus on the design of SU’s optimal OSA policy in the presence of RPUs, and investigate the tradeoff between the SU throughput maximization and the RPU transmission protection. First, we propose a 4-state Markov chain model for the RPU’s channel occupancy (as opposed to the conventional 2-state (on/off) counterpart for the non-reactive case), where the expanded state space and state transition probabilities are used to specify the reactions of the PU with or without the SU transmit collision. Under this model, we formulate the optimal OSA design for the SU’s throughput maximization as a finite-horizon POMDP problem. Due to imperfect sensing, SU’s transmit collision to active PUs is not completely avoidable. As in [2], we adopt the conditional collision probability constraint to protect the PU. We derive the optimal policy for the SU and develop a separation principle to reduce the computational complexity of the optimal policy. By numerical results, we show some interesting observations on the PU and SU throughput tradeoff under the new RPU model.

The rest of this paper is organized as follows. Section II proposes the channel occupancy model for RPUs and describes the corresponding CR network model. Section III formulates the OSA design under the RPU model as a constrained POMDP problem. Section IV studies the optimal policy via applying a separation principle. Numerical results and related discussions are presented in Section V. The paper is concluded.
in Section VI.

II. SYSTEM MODEL

A. Channel Occupancy Model for Reactive Primary User

Fig. 1 gives a typical example of the channel occupancy model designed for NR PUs [1]-[4], which is a 2-state Markov chain with states ‘0’ and ‘1’, denoting whether the channel is busy or idle, respectively. In a time-slotted system, the states change slot by slot according to transition probabilities $\alpha$ and $\beta$ shown in Fig. 1. Unfortunately, this model cannot reflect a RPU’s channel occupancy state (COS). RPUs usually reduce their willingness to access a channel if a collision occurs and become more willing to transmit when the environment becomes better (no collision is observed).

To reflect more practical PU’s behavior in real life, we propose an enhanced COS model. The new model is in general composed of multiple levels, in each of which the PU has a different probability to transmit. For simplicity, two levels are considered in this paper: Level 0 and Level 1. A RPU has higher willingness to occupy a channel when it is in Level 0 than in Level 1. As a result, the enhanced COS model becomes a 4-state Markov chain as shown in Fig. 2, with each level having two states (busy or idle). For convenience, we use 2 bits to represent the total 4 states. The first bit denotes the located level and the second bit denotes whether the channel is busy or idle. For example, state ‘01’ means that the COS locates in Level 0 and the channel is idle.

The proposed COS evolves according to the RPU’s COS in the previous slot and the SU’s action in the current slot. Suppose that there are no SUs in the network and the RPU stays in Level 0, which is exactly the model shown in Fig. 1 with transition probabilities $\alpha_0$ and $\beta_0$. Now suppose the SU exists. If the SU decides not to access the channel, or it accesses and the RPU locates at state ‘01’, no collision will be caused and the RPU will stay in Level 0. However, if the SU accesses the channel and the RPU’s state is ‘00’, the RPU will react to the resulted collision: its COS will transit to Level 1, with probability $\alpha_1$ to state ‘11’ and probability $1 - \alpha_1$ to state ‘10’. We assume $\alpha_1 \geq \alpha_0$ to reflect the reduced willingness of the RPU to transmit. In Level 1, the RPU will stay in this level if the SU continues to access the channel. However, if the SU does not access and the RPU is at state ‘10’, there is no collision and the RPU will perceive that the environment has become better for its transmission. As a result, the RPU will increase its willingness to use the channel and thus return to Level 0, with transition probabilities $\alpha_0$ to state ‘01’ and $1 - \alpha_0$ to state ‘00’, respectively. Since the RPU’s state transition probabilities are related to the SU’s actions when the state is ‘00’ or ‘10’, we denote transition probabilities in terms of the SU’s action in Fig. 2. When the state is ‘01’ or ‘11’, the transition probabilities are not affected by the SU’s action. This is because no collision will occur if the RPU does not occupy the channel. We assume $\beta_1 \geq \beta_0$ to be consistent with the RPU’s different willingness to transmit at each level.

Fig. 1. Channel occupancy model for the non-reactive (NR) PU.

Fig. 2. Channel occupancy model for the reactive PU (RPU).

B. Network Model

Suppose that there are one SU and $N$ PUs in the CR network. Each PU is preassigned one separate channel and the COS of each PU evolves independently according to the RPU model in Fig. 2. We assume synchronized time-slotted transmission for all the PUs and SU as in [1]-[3]. For simplicity, we assume that each channel has unit bandwidth $B_a = 1 \ (\forall a \in \{1, 2, ..., N\})$.

Following similar joint PHY-MAC layer design approach in [2], we consider the SU’s operations as follows: at the beginning of each slot $t$, the SU transmitter first selects a channel $a(t) \in \delta_S$ to sense, where $\delta_S = \{1, 2, ..., N\}$. The SU then access the channel if the sensed result $\Theta_a(t) \in \{0, 1\}$, the SU decides a pair of transmission probabilities $(f_a(0, t), f_a(1, t)) \in [0, 1]^2$ for this channel, where $f_a(\theta, t)$ is the transmission probability on channel $a$ in slot $t$ with $\Theta_a(t) = \theta$. Denoting $\Phi_a(t) \in \{1(\text{access}), 0(\text{not access})\}$ as the access action on channel $a$ in slot $t$, $f_a(\theta, t)$ can then be expressed as $f_a(\theta, t) = \Pr\{\Phi_a(t) = 1|\Theta_a(t) = \theta\}$.

After the transmission, the SU transmitter will receive a feedback information $K_a(t) \in \{0, 1\}$ from the SU receiver. $K_a(t) = 1$ means that the SU’s information is transmitted successfully, while $K_a(t) = 0$ represents that one of the following two events has occurred: 1) SU does not transmit in the current slot; 2) the SU transmits but the SU receiver fails to receive the correct transmitted information because the RPU’s current COS is busy and thus collision occurs.

III. CONSTRAINED POMDP FORMULATION

The sequential decision-making process for the SU in Section II.B can be formulated as a POMDP problem.
A. POMDP Elements

1) State Space: Let the state space \( S = \{ (C, L, T) \} \), where \( C \in \mathcal{A}_S \), \( L \in \{0, 1\} \) and \( T \in \{0, 1\} \) represent a channel, its located level and “busy/idle” state, respectively.

2) Belief Matrix: The SU’s belief in probability (given its past decisions and observations) for the RPU’s initial COS at slot \( t \) over different channels is given by the belief matrix \( \Lambda(t) \), which is defined in (1). For convenience, the 4 states, 00, 01, 10, 11, are labeled as 0, 1, 2, 3, respectively.

\[
\Lambda(t) = \begin{bmatrix}
\lambda_{00}(t) & \lambda_{01}(t) & \lambda_{10}(t) & \lambda_{11}(t) \\
\lambda_{20}(t) & \lambda_{21}(t) & \lambda_{22}(t) & \lambda_{23}(t) \\
\vdots & \vdots & \vdots & \vdots \\
\lambda_{N0}(t) & \lambda_{N1}(t) & \lambda_{N2}(t) & \lambda_{N3}(t)
\end{bmatrix}
\]

(1)

In (1), \( \lambda_{nj}(t) \) represents the conditional probability that channel \( n \) locates at the \( j \)th state at the beginning of slot \( t \) with \( \sum_{j=0}^{3} \lambda_{nj}(t) = 1 \hspace{1em} (\forall n \in \{1, 2, \ldots, N\}) \). \( \Lambda(t) \) is updated slot by slot and the details are given in Section III.B. Similarly to [2] and [9], it is easy to show that \( \Lambda(t) \) is a sufficient statistic for the SU to decide its action at slot \( t \). Assume that the SU knows the initial belief matrix \( \Lambda(1) \) at \( t = 1 \).

3) Policy, Action Space and Observation: Similarly as in [2], there are three policies in our design: (A) A deterministic sensing policy \( \pi_s \), which maps a belief matrix \( \Lambda(t) \) to a channel \( a(t) \in \mathcal{A}_s \) to sense in slot \( t \); (B) A deterministic sensor operating policy \( \pi_s \), which specifies a feasible sensor operating point \( (c_a(t), \delta_a(t)) \in \mathcal{A}_4(a(t)) \) based on \( \Lambda(t) \) and the channel \( a(t) \) selected; (C) A randomized access policy, which determines the transmission probabilities \( (f_a(0, t), f_a(1, t)) \) based on \( \Lambda(t) \) and the sensing result \( \Theta(t) \in \{0, 1\} \). Hence, the composite action space is given by \( \mathcal{A} = \{(a(t), (c_a(t), \delta_a(t)), (f_a(0, t), f_a(1, t)))\} \). Also as in [2], we use the SU feedback information \( K_a(t) \in \{0, 1\} \) as the observation of the POMDP.

Fig. 3. State Transition Probabilities in the COS model for the RPU.

4) Transition Probabilities: Denote any two arbitrary states of this POMDP as \( s_1 = \{C_1, L_1, T_1\} \in S \) and \( s_2 = \{C_2, L_2, T_2\} \in S \). If \( C_1 \neq C_2 \), \( s_1 \) and \( s_2 \) are locating at two different channels. The state transition probability between \( s_1 \) and \( s_2 \) is thus 0 because these two channels evolve independently. If \( C_1 = C_2 \), their transition probability is reduced to the COS transition probability, which is related to the SU’s action over this particular channel as shown in Fig. 2. Denote \( P_a(i|j, A) \) as the COS transition probability from state \( j \in \{0, 1, 2, 3\} \) to \( i \in \{0, 1, 2, 3\} \) over channel \( n \hspace{1em} (\forall n \in \{1, 2, \ldots, N\}) \) with the SU’s action \( A = (a(t), a, (c_a(t), \delta_a(t)), (f_a(0, t), f_a(1, t))) \in \mathcal{A} \). If \( n \neq a \), the SU does not select channel \( n \) for sensing. Fig. 3(a) shows the COS transition probabilities under this case, which can be easily obtained from Fig. 2. If \( n = a \), channel \( a \) is selected for sensing and (probably to access). Fig. 3(b) shows the COS transition probabilities under this case: if the COS is ‘00’ or ‘10’, the transition probabilities are independent of the SU’s action, as shown in Fig. 3(a); otherwise, according to Fig. 2, they are subject to the SU’s action \( \Phi_a(t) \). We use \( P_a(i|j, \Phi_a(t)) \) to denote the COS transition probability over channel \( a \) in Fig. 2 from state \( j \in \{0, 2\} \) to \( i \in \{0, 1, 2, 3\} \) with action access \( \Phi_a(t) \). Let \( \Phi_a(t) = \phi \) and the observation \( \Theta_a(t) = \theta \). Then we obtain \( P_a(i|j, A) \) in Fig. 3(b).

\[
P_a(i|j, A) = \sum_{\phi=0}^{1} P_a(\phi|j, A) \times P_a(i|\phi, j, A) = \sum_{\theta=0}^{1} P_a(\theta|j) \times P_a(\phi|\theta) \times P_a(i|j, \phi).
\]

(3)

From Fig. 2, the transition probabilities starting from states ‘00’ and ‘10’ are the same. Thus, \( P_a(i|j = 2, A) = P_a(i|j = 0, A) \), \( \forall i \in \{0, 1, 2, 3\} \). By applying (3), the transition probabilities are given as:

\[
\begin{align*}
P_a(i = 0|j = 0, A) &= (1 - \mu_a(0))(1 - \alpha_0) \\
P_a(i = 1|j = 0, A) &= (1 - \mu_a(0))\alpha_0 \\
P_a(i = 2|j = 0, A) &= \mu_a(0)(1 - \alpha_1) \\
P_a(i = 3|j = 0, A) &= \mu_a(0)\alpha_1
\end{align*}
\]

where \( \mu_a(0) = (1 - \delta_a(0))f_a(0, t) + \delta_a(0)f_a(1, t) \).

5) Immediate Reward: If the SU transmits successfully in slot \( t \), it will obtain one unit throughput at the end of the slot. Define the unit throughput as its immediate reward \( R(t) \). Let \( I(a, t) = 0 \) and \( I(a, t) = 1 \) represent channel \( a \) is busy or idle in slot \( t \), respectively. \( R(t) \) is then expressed as:

\[
R(t) = K_a(t)B_a = \Phi_a(t)I(a, t)B_a.
\]

(4)

B. Problem Formulation

Our goal is to maximize the SU’s expected total throughput in \( T \) slots while protecting the PU under the RPU model by choosing the best sensing policy \( \pi_s \), the best sensor operating policy \( \pi_s^i \) and the best access policy \( \pi_s^e \) for the SU. We adopt the instantaneous conditional collision probability constraint as the method to protect the RPU [2], where the conditional collision probability \( P_c(t) \) is defined as \( P_c(t) = \Pr[\Phi_a(t) = 1|I(a, t) = 0] \). The optimization problem is thus given by:

\[
\{\pi_s^i, \pi_s^e, \pi_a^e\} = \arg\max_{\pi_s, \pi_s^i, \pi_s^e} E_{\pi_s, \pi_s^i, \pi_s^e}\{\sum_{t=1}^{T} R(t)|\Lambda(1)\}.
\]

s.t.

\[
P_a(t) \leq \zeta, \hspace{1em} \forall t \in \{1, \ldots, T\}.
\]

(6)

where \( \zeta \) is the collision probability constraint. By a similar derivation as in [2], \( P_a(t) \) can be further specified as:

\[
P_a(t) = (1 - \delta_a(t))f_a(0, t) + \delta_a(t)f_a(1, t) \leq \zeta.
\]

(7)
The objective function in (5) can be decoupled into T subproblems without the loss of optimality [10]. Each subproblem maximizes a value function $V_t(\Lambda(t))$ ($\forall t \in \{1, 2, ..., T\}$):

$$V_t(\Lambda(t)) = \max_{A \in \mathcal{K}} \sum_{\substack{i, j = 0 \to \infty \atop k = 0 \to \infty}} \sum_{\substack{i, j = 0 \to \infty \atop k = 0 \to \infty}} \lambda_{ai}(t)P_a(j|i, A)U_A(k|j) \times [kB_a + V_{t+1}(T(\Lambda(t)|A, k))], \quad 1 \leq t \leq T - 1 \quad (8)$$

The above separation principle is analogous to that in [2], but it requires a different proof. Due to the space constraint, we provide a sketch of the proof in the following for the separation principle under the RPU model; the complete proof is left to the journal version of this paper [11].

Proof of the separation principle under the RPU model: Based on the mathematical induction and by computing the value functions backward in time, we first show that in any slot $t$ ($1 \leq t \leq T$) for any selected channel $a(t) = a$ ($\forall a \in \{1, 2, ..., N\}$), $V_t(\Lambda(t))$ can be represented in the form that $V_t(\Lambda(t)) = D(\lambda_{ai}(t) + \lambda_{aj}(t)) + H$ where $D$, $F$, $H$ are non-negative constants and satisfy $H \geq F$. Then, we show that with the given future reward $V_{t+1}(A(t+1))$ represented in the above form, the optimal operating policy and the optimal access policy are independent of the optimal sensing policy. Hence, we can separate the OSA design problem into the two steps shown above, without the loss of optimality.

V. NUMERICAL RESULTS

We consider an energy detector based spectrum sensor for the SU, where the background noise and the received PU signal are modeled as independent white Gaussian processes. Let $M$ be the number of PU signal measurements, and $\eta_n$ be the decision threshold. Let $\sigma_{n,0}^2$ and $\sigma_{n,1}^2$ denote the power of the noise and received PU’s signal at channel $n$, respectively. Under the Neyman-Pearson (NP) criterion, the PFA and the PM are obtained as [12]: $\delta_n = \gamma(\frac{M}{F}, \frac{\eta_n}{2\sigma_{n,0}^2 + \sigma_{n,1}^2})$ and $\epsilon_n = 1 - \gamma(\frac{M}{F}, \frac{\eta_n}{2\sigma_{n,0}^2})$, respectively, where $\gamma(a, m) = (1/F(m)) \int_0^m t^{a-1} e^{-t} dt$ is the incomplete gamma function. The optimal decision threshold $\eta_n$ of the energy detector is chosen such that $\delta_n = \zeta$. We set $\sigma_{n,0}^2 = 0$ dB, $\sigma_{n,1}^2 = 5$ dB ($\forall n \in \{1, ..., N\}$), $M = 30$, and $\zeta = 0.05$.

We investigate the tradeoff between the SU’s and the PU’s throughput by simulation. We consider three scenarios: 1) under the NR PU model shown in Fig. 1; 2) under the RPU model shown in Fig. 2; 3) for the case of model mismatch, which occurs when the SU implements the optimal access policies assuming the NR PU model, but actually under the RPU model. We compare the throughput in these three scenarios under each of the following 2 cases. Case 1: $N = 3$; $\alpha_0 = 0.3$, $\beta_0 = 0.1$, $\alpha_1 = 0.8$, and $\beta_1 = 0.8$ for the RPU and $\alpha_0 = 0.5$, $\beta_0 = 0.1$, $\alpha_1 = 0.5$, $\beta_1 = 0.8$ for the NR PU. Case 2: $N = 3$; all the channels adopt the same $(\alpha_0, \beta_0, \alpha_1, \beta_1)$ and $(\alpha, \beta)$ as
those in Case 1 for the RPU and the NR PU, respectively. The PU’s initial access probability to a channel is set to be the stationary distribution of the underlying Markov Chain, which is computed as \(\frac{1-\alpha}{1-\alpha - \beta_0}\) and \(\frac{1-\beta}{1-\alpha - \beta}\) for the RPU and the NR PU, respectively; clearly, they are same with \((\alpha, \beta) = (\alpha_0, \beta_0)\).

We adopt the normalized expected throughput as the throughput measure of the SU and PU. For SU, the normalized expected throughput is defined as \(V_1(\Lambda(1))/T\). For PU, suppose the SU selects channel \(a\) in slot \(t\), the PU’s expected immediate throughput can be expressed as \(E\{R_P(t)\} = \sum_{n=1}^{N} \sum_{i}^{1} \sum_{j}^{1} \lambda_{nj}(t) P_n(j|j, A) \times B_n + \sum_{n=0}^{N} \sum_{j}^{1} \lambda_{nj}(t) P_n(j|i, A) \times (1 - \delta^*_n) \times B_{\delta^*}\). Thus, the PU’s normalized expected total throughput in \(T\) slots is \((1/T) \times E\{\sum_{t=1}^{T} R_P(t)\}\). For SU, the normalized throughput is \(E{\{\sum_{t=1}^{T} \pi^*_n, \pi^*_\delta, \pi^*_c}\}\).

First, we compare the SU’s throughput under different cases in Fig. 4 using the POMDP solver software [13]. By comparing the SU’s throughput under the NR PU model and the RPU model in the two cases, we observe that the SU achieves much throughput in the latter than in the former model. This is because the RPU reduces but the NR PU does not change the channel access probability after a collision with the SU occurs. Since the NR PU has the same channel access probability, when \(N = 1\), the expected channel access opportunities are unchanged over time for the SU. As a result, the SU’s throughput is a constant in Case 1 under the NR PU model. In addition, we observe that when \(N = 1\) for the case of model mismatch, the SU obtains the same throughput as that under the RPU model, i.e., model mismatch does not affect the SU’s throughput. The reason for it is that the optimal policies \(\pi^*_n\) and \(\pi^*_\delta\) can be shown to be the same under both RPU and NR PU models; thus, since \(N = 1\), the optimal policy \(\pi^*_c\) is also the same. However, this is not the case when \(N > 1\). For example, in the model mismatch of Case 2, the throughput loss occurs when \(T \geq 3\).

Next, we study the PU’s throughput. Fig. 5 shows the PU’s throughput under the SU’s optimal policies for both NR PU and RPU. We observe that the PU achieves more throughput in the former than in the later model in both cases. The PU’s throughput remains unchanged over time in Case 1 and slightly increases in Case 2 for NR PU; thus, the NR PU is protected as expected. Hence, taking the NR PU’s throughput as reference, the RPU’s throughput is not well protected under the same conditional collision probability constraint. This is because under the RPU model, when collision with the SU occurs, the PU will transit from Level 0 to Level 1 with reduced channel access probability. However, with the conditional collision probability constraint, the SU’s access probability when the PU is busy in Level 1 remains the same as that in Level 0. Thus, the PU is less likely to return back to Level 0. Hence, the conditional collision probability constraint proposed to protect the NR PU can be ineffective in the RPU case.

VI. CONCLUSION

In this paper, we propose a new transmission model for the RPU, where the probability for the PU to access a channel is related to the SU’s access policy. Under this model, we formulate the optimal OSA design problem for the SU as a POMDP. Moreover, we develop a separation principle to obtain the optimal access policies with implementable complexity. In addition, we show a new tradeoff between the SU’s throughput maximization and the PU transmission protection under the RPU model, which reveals that the conditional collision probability constraint proposed to protect the NR PU can be ineffective in the RPU case. In the future work, we will investigate new criteria to more effectively protect the PU transmission for the SU OSA under the RPU model.

REFERENCES